

A general product of tensors with applications [☆]

Jia-Yu Shao^{a,*}

^a*Department of Mathematics, Tongji University, Shanghai 200092, China*

Abstract

We define a general product of two n -dimensional tensors \mathbb{A} and \mathbb{B} with orders $m \geq 2$ and $k \geq 1$, respectively. This product is a generalization of the usual matrix product, and satisfies the associative law. Using this product, many concepts and known results of tensors can be simply expressed and/or proved, and a number of applications of this product will be given. Using this tensor product and some properties on the resultant of a system of homogeneous equations on n variables, we define the similarity and congruence of tensors (which are also the generalizations of the corresponding relations for matrices), and prove that similar tensors have the same characteristic polynomials. We study two special kinds of similarity: permutational similarity and diagonal similarity, and their applications in the study of the spectra of hypergraphs and nonnegative irreducible tensors. We define the direct product of tensors (in matrix case it is also called the Kronecker product), and give its applications in the study of the spectra of two kinds of the products of hypergraphs. We also give applications of this tensor product in the study of nonnegative tensors, including a characterization of primitive tensors, the upper bounds of primitive degrees and the cyclic indices of some nonnegative irreducible tensors.

AMS classification: 15A18; 15A69; 05C50

Keywords: tensor, product, eigenvalue, spectrum, characteristic polynomial, hypergraph;

1. Introduction

In recent years, the study of tensors and the spectra of tensors (and hypergraphs) with their various applications has attracted much attention and interest, since the work of L.Qi ([12]) and L.H.Lim ([8]) in 2005.

As is in [12], an order m dimension n tensor $\mathbb{A} = (a_{i_1 i_2 \dots i_m})_{1 \leq i_j \leq n \ (j=1, \dots, m)}$ over the complex field \mathbb{C} is a multidimensional array with all entries $a_{i_1 i_2 \dots i_m} \in \mathbb{C}$ ($i_1, \dots, i_m \in [n] = \{1, \dots, n\}$). For a vector $x = (x_1, \dots, x_n)^T \in \mathbb{C}^n$, let $\mathbb{A}x^{m-1}$ be a vector in \mathbb{C}^n whose i th component is defined as the following:

$$(\mathbb{A}x^{m-1})_i = \sum_{i_2, \dots, i_m=1}^n a_{i i_2 \dots i_m} x_{i_2} \cdots x_{i_m} \quad (1.1)$$

and let $x^{[r]} = (x_1^r, \dots, x_n^r)^T$. Then ([3,12]) a number $\lambda \in \mathbb{C}$ is called an eigenvalue of \mathbb{A} if there exists a nonzero vector $x \in \mathbb{C}^n$ such that

$$\mathbb{A}x^{m-1} = \lambda x^{[m-1]} \quad (1.2)$$

[☆]Research supported by the National Science Foundation of China No.11231004

*Corresponding author.

Email address: jyshao@tongji.edu.cn (Jia-Yu Shao)

and in this case, x is called an eigenvector of \mathbb{A} corresponding to the eigenvalue λ . L.Qi also defined several other types of eigenvalues (and eigenvectors) in [12].

Since L.Qi introduced the definitions of various types of the eigenvalues (and eigenvectors) of tensors, quite many interesting research works in this area appeared (see [2-5] and [8-15]). But noticing that, the term $\mathbb{A}x^{m-1}$ in the left side of (1.1) is only a special notation (its inside part x^{m-1} does not have specific meaning), and also we can not view the notation $\mathbb{A}x^{m-1}$ as some general operation between \mathbb{A} and x^{m-1} . Anyway, we have not seen some kinds of general product between tensors up to now.

In this paper, we will introduce (in Definition 1.1) a general tensor product defined for any two tensors of dimension n (one has order $m \geq 2$ and another one has order $k \geq 1$). We will see that, the product of two dimension n tensors is still a tensor of dimension n (but in general will have different order). This tensor product also possess many properties. Among them the most important and useful property is the associative law (see Theorem 1.1).

Besides these properties, we can also see that, when the two tensors are matrices (or the first one is a matrix and the second one is a vector), our definition of tensor product coincides with the usual matrix product, so it is a generalization of the matrix product.

In this paper, we will also give a number of applications of this tensor product. we will see that, by using this tensor product, not only many mathematical expressions can be simplified (for example, the term $\mathbb{A}x^{m-1}$ in the left side of (1.1) can now be directly simplified as $\mathbb{A}x$), and many known results can be simply stated and proved, but we can also obtain many new results (see sections 2,3 and 4).

In section 2, we first use this tensor product and some properties on resultant (namely, the hyperdeterminant) to define the similarity of tensors. We prove that the similar tensors have the same characteristic polynomials, thus have the same spectrum (as multiset). We also study two special kinds of similarity: permutation similarity and diagonal similarity. Using the permutation similarity we can show that the spectrum of a hypergraph is independent of the ordering of its vertices (but the adjacency tensor of a hypergraph depends on the ordering of its vertices). Using the diagonal similarity, we obtain a result on the symmetry of the spectrum of a nonnegative irreducible tensor which is stronger than the result in [14]. That is: if a nonnegative irreducible tensor \mathbb{A} has cyclic index k , then the whole spectrum of \mathbb{A} is invariant under a rotation of angle $2\pi/k$ in the complex plane. We also define in section 2 the congruence relation between tensors and apply this relation to the study of \mathbf{E} -eigenvalues and \mathbf{E} -eigenvectors ([12]).

In section 3, we generalize the direct product of matrices to the direct product of tensors (of the same order, but may be different dimensions). We first prove a relation (Theorem 3.1) between the direct product of tensors and the general product of tensors defined in Definition 1.1. Then we use this to study the spectra of two kinds of products of hypergraphs.

In section 4, we apply this tensor product to the study of nonnegative tensors. We give a simple characterization for the primitivity of a nonnegative tensor in terms of the powers of \mathbb{A} . We show that the primitivity of a nonnegative tensor \mathbb{A} is uniquely determined by the zero pattern of \mathbb{A} . From this we show the boundedness of the primitive degrees (for fixed n and m), and give some upper bounds for the primitive degrees of the primitive tensors. We also propose a conjecture concerning the upper bound of the primitive degrees.

For the sake of simplicity, we sometime use the following "binary notation" for the subscripts of the tensor. Namely, we will write $a_{i_1 i_2 \dots i_m}$ as $a_{i_1 \alpha}$, where $\alpha = i_2 \dots i_m \in [n]^{m-1}$.

Now we define the general product of two n -dimensional tensors.

Definition 1.1: Let \mathbb{A} (and \mathbb{B}) be an order $m \geq 2$ (and order $k \geq 1$), dimension n tensor, respectively. Define the product $\mathbb{A}\mathbb{B}$ to be the following tensor \mathbb{C} of order $(m-1)(k-1)+1$ and dimension n :

$$c_{i\alpha_1 \cdots \alpha_{m-1}} = \sum_{i_2, \dots, i_m=1}^n a_{ii_2 \cdots i_m} b_{i_2 \alpha_1} \cdots b_{i_m \alpha_{m-1}} \quad (i \in [n], \alpha_1, \dots, \alpha_{m-1} \in [n]^{k-1})$$

Example 1.1: When $k = 1$ and $\mathbb{B} = x \in \mathbb{C}^n$ is a vector of dimension n , then $(m-1)(k-1)+1 = 1$. Thus $\mathbb{A}\mathbb{B} = \mathbb{A}x$ is still a vector of dimension n , and we have

$$(\mathbb{A}x)_i = (\mathbb{A}\mathbb{B})_i = c_i = \sum_{i_2, \dots, i_m=1}^n a_{ii_2 \cdots i_m} x_{i_2} \cdots x_{i_m} = (\mathbb{A}x^{m-1})_i \quad (\text{where } \mathbb{A}x^{m-1} \text{ is defined by (1.2)})$$

Thus we have $\mathbb{A}x^{m-1} = \mathbb{A}x$. So the first application of the tensor product defined above is that now $\mathbb{A}x^{m-1}$ can be simply written as $\mathbb{A}x$.

Example 1.2: (1). If $k = 2$ (namely \mathbb{B} is a matrix of order n), then $(m-1)(k-1)+1 = m$. So in this case $\mathbb{A}\mathbb{B}$ is still a tensor of order m .

(2). If $m = 2$ (namely \mathbb{A} is a matrix of order n), then $(m-1)(k-1)+1 = k$. So in this case $\mathbb{A}\mathbb{B}$ is still a tensor of order k .

(3). It is easy to check from the definition that $I_n \mathbb{A} = \mathbb{A} = \mathbb{A} I_n$, where I_n is the identity matrix of order n .

Example 1.3: In case when both \mathbb{A} and \mathbb{B} are matrices, or when \mathbb{A} is a matrix and \mathbb{B} is a vector, our tensor product $\mathbb{A}\mathbb{B}$ coincides with the usual matrix product. So it is a generalization of the matrix product.

Proposition 1.1: The tensor product defined in Definition 1.1 has the following properties.

- (1). (The left distributivity): $(\mathbb{A}_1 + \mathbb{A}_2)\mathbb{B} = \mathbb{A}_1\mathbb{B} + \mathbb{A}_2\mathbb{B}$.
- (2). (The right distributivity in the case of $m = 2$): $\mathbb{A}(\mathbb{B}_1 + \mathbb{B}_2) = \mathbb{A}\mathbb{B}_1 + \mathbb{A}\mathbb{B}_2$ (when \mathbb{A} is a matrix).
- (Note that in general when \mathbb{A} is not a matrix, then the right distributivity does not hold.)
- (3). $(\lambda \mathbb{A})\mathbb{B} = \lambda(\mathbb{A}\mathbb{B})$. $(\lambda \in \mathbb{C})$
- (4). $\mathbb{A}(\lambda \mathbb{B}) = \lambda^{m-1}(\mathbb{A}\mathbb{B})$. $(\lambda \in \mathbb{C})$

Next we will show that the associative law of this tensor product holds. The following identity will be needed in the proof of the associative law.

$$\prod_{j=1}^m \sum_{t_{j1}, \dots, t_{jk}=1}^n f(j, t_{j1}, \dots, t_{jk}) = \sum_{t_{jh}=1(1 \leq j \leq m; 1 \leq h \leq k)}^n \prod_{j=1}^m f(j, t_{j1}, \dots, t_{jk}) \quad (1.1)$$

For the simplicity of the notation, we assume in the following that the tensors \mathbb{A} , \mathbb{B} and \mathbb{C} have the orders $m+1$, $k+1$ and $r+1$, respectively.

Theorem 1.1 (The associative law of the tensor product): Let \mathbb{A} (and \mathbb{B} , \mathbb{C}) be an order $m+1$ (and order $k+1$, order $r+1$), dimension n tensor, respectively. Then we have

$$\mathbb{A}(\mathbb{B}\mathbb{C}) = (\mathbb{A}\mathbb{B})\mathbb{C}$$

Proof. For $\beta_1, \dots, \beta_m \in ([n]^r)^k$, we write:

$$\beta_1 = \theta_{11} \cdots \theta_{1k}, \quad \dots, \quad \beta_m = \theta_{m1} \cdots \theta_{mk} \quad (\theta_{ij} \in [n]^r, \quad i = 1, \dots, m; \quad j = 1, \dots, k.)$$

Then we have:

$$\begin{aligned}
(\mathbb{A}(\mathbb{B}\mathbb{C}))_{i\beta_1\cdots\beta_m} &= \sum_{i_1,\cdots,i_m=1}^n a_{ii_1\cdots i_m} \left(\prod_{j=1}^m (\mathbb{B}\mathbb{C})_{i_j\beta_j} \right) \\
&= \sum_{i_1,\cdots,i_m=1}^n a_{ii_1\cdots i_m} \left(\prod_{j=1}^m (\mathbb{B}\mathbb{C})_{i_j\theta_{j1}\cdots\theta_{jk}} \right) \\
&= \sum_{i_1,\cdots,i_m=1}^n a_{ii_1\cdots i_m} \left(\prod_{j=1}^m \sum_{t_{j1},\cdots,t_{jk}=1}^n b_{i_j t_{j1}\cdots t_{jk}} (c_{t_{j1}\theta_{j1}} \cdots c_{t_{jk}\theta_{jk}}) \right) \\
&\stackrel{(*)}{=} \sum_{i_1,\cdots,i_m=1}^n a_{ii_1\cdots i_m} \sum_{t_{jh}=1(1\leq j\leq m; 1\leq h\leq k)}^n \left(\prod_{j=1}^m b_{i_j t_{j1}\cdots t_{jk}} (c_{t_{j1}\theta_{j1}} \cdots c_{t_{jk}\theta_{jk}}) \right)
\end{aligned} \tag{1.2}$$

where the equation (*) follows from equation (1.1).

On the other hand, for $\alpha_1, \cdots, \alpha_m \in [n]^k$, we write:

$$\alpha_1 = t_{11} \cdots t_{1k}, \quad \cdots, \quad \alpha_m = t_{m1} \cdots t_{mk} \quad (t_{ij} \in [n], \quad i = 1, \cdots, m; \quad j = 1, \cdots, k.)$$

Then we also have:

$$\begin{aligned}
((\mathbb{A}\mathbb{B})\mathbb{C})_{i\beta_1\cdots\beta_m} &= \sum_{\alpha_1,\cdots,\alpha_m \in [n]^k} (\mathbb{A}\mathbb{B})_{i\alpha_1\cdots\alpha_m} \left(\prod_{j=1}^m (c_{t_{j1}\theta_{j1}} \cdots c_{t_{jk}\theta_{jk}}) \right) \\
&= \sum_{t_{jh}=1(1\leq j\leq m; 1\leq h\leq k)}^n \sum_{i_1,\cdots,i_m=1}^n a_{ii_1\cdots i_m} \left(\prod_{j=1}^m b_{i_j\alpha_j} \right) \left(\prod_{j=1}^m (c_{t_{j1}\theta_{j1}} \cdots c_{t_{jk}\theta_{jk}}) \right) \\
&= \sum_{i_1,\cdots,i_m=1}^n a_{ii_1\cdots i_m} \sum_{t_{jh}=1(1\leq j\leq m; 1\leq h\leq k)}^n \left(\prod_{j=1}^m b_{i_j t_{j1}\cdots t_{jk}} (c_{t_{j1}\theta_{j1}} \cdots c_{t_{jk}\theta_{jk}}) \right)
\end{aligned} \tag{1.3}$$

Comparing the right hand sides of (1.2) and (1.3), we obtain $\mathbb{A}(\mathbb{B}\mathbb{C}) = (\mathbb{A}\mathbb{B})\mathbb{C}$. \square

By the associative law, we can henceforth write $(\mathbb{A}\mathbb{B})\mathbb{C}$ and $\mathbb{A}(\mathbb{B}\mathbb{C})$ as $\mathbb{A}\mathbb{B}\mathbb{C}$. We can also define \mathbb{A}^k as the product of k many tensors \mathbb{A} . Furthermore, by the left distributive law and right distributive law for matrices, we have:

$$P(\mathbb{A} + \mathbb{B})Q = P\mathbb{A}Q + P\mathbb{B}Q \quad (\text{when } P, Q \text{ are both matrices}) \tag{1.4}$$

The unit tensor of order m and dimension n is the tensor $\mathbb{I} = (\delta_{i_1,i_2,\cdots,i_m})$ with entries as follows:

$$\delta_{i_1,i_2,\cdots,i_m} = \begin{cases} 1 & \text{if } i_1 = i_2 = \cdots = i_m \\ 0 & \text{otherwise} \end{cases}$$

It is easy to see from the definition that $\mathbb{I}x = x^{[m-1]}$. Thus equation (1.2) can be rewritten as

$$(\lambda\mathbb{I} - \mathbb{A})x = 0 \tag{1.5}$$

2. The similarity and congruence of tensors

In this section, we first use the tensor product defined in section 1, together with some properties on the resultant of a system of homogeneous equations on n variables, to define the similarity of tensors (which is a generalization of the similarity of matrices), and then prove that the similar tensors have the same characteristic polynomials, hence the same spectrum (as a multiset). We further study two special kinds of similarity: permutation similarity and diagonal similarity, then apply them respectively in the study of the spectra of hypergraphs, and in the study of the symmetry of the spectra of nonnegative irreducible tensors. We also define the congruence relation between tensors, and show how this relation can be applied in the study of \mathbf{E} -eigenvalues and \mathbf{E} -eigenvectors of tensors ([12]).

In [6], D.Cox et al used the resultant to define the hyperdeterminant $\det(\mathbb{A})$ of a tensor \mathbb{A} .

Definition 2.1 ([6], hyperdeterminant): Let \mathbb{A} be an order m dimension n tensor. Then the hyperdeterminant of \mathbb{A} , denoted by $\det(\mathbb{A})$, is the resultant of the system of homogeneous equations $\mathbb{A}x = 0$, which is also the unique polynomial on the entries of \mathbb{A} (viewed as independent variables) satisfying the following three conditions:

- (1) $\det(\mathbb{A}) = 0$ if and only if the system of homogeneous equations $\mathbb{A}x = 0$ has a nonzero solution.
- (2) $\det(\mathbb{I}) = 1$, where \mathbb{I} is the unit tensor.
- (3) $\det(\mathbb{A})$ is an irreducible polynomial on the entries of \mathbb{A} , when the entries of \mathbb{A} are viewed as independent variables.

The proof of the existence and uniqueness of the polynomial satisfying the above three conditions can be found in [7].

Definition 2.2: Let \mathbb{A} be an order m dimension n tensor. Then the characteristic polynomial of \mathbb{A} , denoted by $\phi_{\mathbb{A}}(\lambda)$, is the hyperdeterminant $\det(\lambda\mathbb{I} - \mathbb{A})$.

It is easy to see from the above discussions and the equation (1.5) that, λ_0 is an eigenvalue of \mathbb{A} if and only if it is a root of the characteristic polynomial of \mathbb{A} .

In the following study of the similarity and congruence of tensors, we need to use the following formula for the entries of the tensor $P\mathbb{A}Q$ (where P, Q are both matrices).

$$(P\mathbb{A}Q)_{i_1 \dots i_m} = \sum_{j_1, \dots, j_m=1}^n a_{j_1 \dots j_m} p_{i_1 j_1} q_{j_2 i_2} \dots q_{j_m i_m} \quad (2.1)$$

By using the three conditions of the hyperdeterminant in Definition 2.1, we can prove the following lemma.

Lemma 2.1: Let \mathbb{A} be an order m dimension n tensor, \mathbb{I} be the order m dimension n unit tensor, P and Q are two matrices of order n . Then we have:

$$\det(P\mathbb{A}Q) = \det(P\mathbb{I}Q)\det(\mathbb{A}) \quad (2.2)$$

Proof. We consider the following two cases.

Case 1: Both P, Q are invertible.

Then by the associative law of the tensor product, we can see that $\mathbb{A}x = 0$ has a nonzero solution x if and only if $(P\mathbb{A}Q)y = 0$ has a nonzero solution y , which implies that $\det(\mathbb{A}) = 0$ if and only if $\det(P\mathbb{A}Q) = 0$.

Now by the definition, $\det(\mathbb{A})$ is irreducible, so $\det(\mathbb{A})$ is a factor of $\det(P\mathbb{A}Q)$. Similarly, $\det(P\mathbb{A}Q)$ is a factor of $\det(\mathbb{A})$ since P, Q are both invertible. Thus there exists some constant $c = c(P, Q)$ (independent of \mathbb{A}) such that $\det(P\mathbb{A}Q) = c \cdot \det(\mathbb{A})$.

Since c is independent of \mathbb{A} , we can take $\mathbb{A} = \mathbb{I}$, which gives us that $c = \det(P\mathbb{I}Q)$, so (2.2) holds in this case.

Case 2: The general case.

Take $P_\epsilon = P + \epsilon I$ and $Q_\epsilon = Q + \epsilon I$. Then P_ϵ and Q_ϵ can be both invertible when ϵ is in some small open interval $(0, \delta)$. Thus by Case 1 we have

$$\det(P_\epsilon \mathbb{A} Q_\epsilon) = \det(P_\epsilon \mathbb{I} Q_\epsilon) \det(\mathbb{A}) \quad (\epsilon \in (0, \delta)) \quad (2.3)$$

Take the limit $\epsilon \rightarrow 0$ on both sides of (2.3), we obtain (2.2). \square

Definition 2.3: Let \mathbb{A} and \mathbb{B} be two order m dimension n tensors. Suppose that there exist two matrices P and Q of order n with $P\mathbb{I}Q = \mathbb{I}$ such that $\mathbb{B} = P\mathbb{A}Q$, then we say that the two tensors \mathbb{A} and \mathbb{B} are similar.

Remark 2.1: If P and Q are two matrices of order n with $P\mathbb{I}Q = \mathbb{I}$, where \mathbb{I} is the order m dimension n unit tensor, then both P and Q are invertible matrices.

Proof. (1). If Q is not invertible, then there exists some nonzero vector $x \in \mathbb{C}^n$ such that $Qx = 0$. Thus by the associative law of the tensor product we have $x^{[m-1]} = \mathbb{I}x = P\mathbb{I}Qx = 0$, which implies that $x = 0$, a contradiction.

(2). If P is not invertible, then there exists some nonzero vector $y \in \mathbb{C}^n$ such that $Py = 0$. Now take $x = Q^{-1}y^{[\frac{1}{m-1}]}$, then x is a nonzero vector since y is. But then we would have $x^{[m-1]} = \mathbb{I}x = P\mathbb{I}Qx = P\mathbb{I}y^{[\frac{1}{m-1}]} = Py = 0$, which implies that $x = 0$, a contradiction. \square

Remark 2.2: The similarity relation of tensors is an equivalent relation.

Proof. The reflection follows from $I_n \mathbb{A} I_n = \mathbb{A}$ (where I_n is the identity matrix). The symmetry follows from Remark 2.1 and the associative law of the tensor product. The transitivity also follows from the associative law of the tensor product.

The following theorem shows that similar tensors have the same characteristic polynomials, and thus they have the same spectrum (as a multiset).

Theorem 2.1: Let \mathbb{A} and \mathbb{B} be two order m dimension n similar tensors. Then we have:

(1). $\det(\mathbb{B}) = \det(\mathbb{A})$.

(2). $\phi_{\mathbb{B}}(\lambda) = \phi_{\mathbb{A}}(\lambda)$.

Proof. By definition, there exist two matrices P and Q of order n with $P\mathbb{I}Q = \mathbb{I}$ and $\mathbb{B} = P\mathbb{A}Q$. Then (1) follows directly from Lemma 2.1.

For (2), by using the equation (1.4) and the relations $P\mathbb{I}Q = \mathbb{I}$ and $\mathbb{B} = P\mathbb{A}Q$, we have

$$P(\lambda\mathbb{I} - \mathbb{A})Q = \lambda P\mathbb{I}Q - P\mathbb{A}Q = \lambda\mathbb{I} - \mathbb{B}$$

Thus $\lambda\mathbb{I} - \mathbb{A}$ and $\lambda\mathbb{I} - \mathbb{B}$ are also similar. Using this and the result (1) of this theorem we have

$$\phi_{\mathbb{B}}(\lambda) = \det(\lambda\mathbb{I} - \mathbb{B}) = \det(\lambda\mathbb{I} - \mathbb{A}) = \phi_{\mathbb{A}}(\lambda).$$

\square

In the following, we consider two special kinds of the similarity of tensors. The first one is the permutation similarity of tensors, and the second one is the diagonal similarity of tensors (also see [14,15]).

Lemma 2.2: Let $\sigma \in S_n$ be a permutation on the set $[n]$, $P = P_\sigma = (p_{ij})$ be the corresponding permutation matrix of σ (where $p_{ij} = 1 \iff j = \sigma(i)$). Let \mathbb{A} and \mathbb{B} be two order m dimension n tensors with $\mathbb{B} = P\mathbb{A}P^T$. Then we have:

- (1). $b_{i_1 \dots i_m} = a_{\sigma(i_1) \dots \sigma(i_m)}$.
- (2). $P\mathbb{I}P^T = \mathbb{I}$. (Thus \mathbb{A} and \mathbb{B} are similar.)

Proof. (1) By using the formula (2.1) we have:

$$b_{i_1 \dots i_m} = \sum_{j_1, \dots, j_m=1}^n a_{j_1 \dots j_m} p_{i_1 j_1} ((P^T)_{j_2 i_2} \dots (P^T)_{j_m i_m}) = \sum_{j_1, \dots, j_m=1}^n a_{j_1 \dots j_m} p_{i_1 j_1} p_{i_2 j_2} \dots p_{i_m j_m} = a_{\sigma(i_1) \dots \sigma(i_m)}$$

(2). From (1) we see that the diagonal elements of \mathbb{B} are those diagonal elements of \mathbb{A} , and the non-diagonal elements of \mathbb{B} are those non-diagonal elements of \mathbb{A} . Thus we have $\mathbb{B} = \mathbb{I}$ when $\mathbb{A} = \mathbb{I}$. Namely $P\mathbb{I}P^T = \mathbb{I}$. \square

From (2) of Lemma 2.2, we see that if $P = P_\sigma$ is a permutation matrix, \mathbb{A} and \mathbb{B} are two order m dimension n tensors with $\mathbb{B} = P\mathbb{A}P^T$, then \mathbb{A} and \mathbb{B} are similar tensors. In this case we say that \mathbb{A} and \mathbb{B} are **permutational similar**.

It is easy to see that if \mathbb{A} and \mathbb{B} are **permutational similar**, and \mathbb{A} is supersymmetric, then \mathbb{B} is also supersymmetric.

An important application of the permutational similar tensors is on the spectra of hypergraphs. According to the definition, the spectrum of a hypergraph is the spectrum of its adjacency tensor. But the adjacency tensor of a hypergraph depends on the ordering of its vertices. Thus a natural question arises: Is the spectrum of a hypergraph independent of the ordering of its vertices?

By using the permutational similar tensors, it is easy to see that if we have two different orderings of the vertices of a hypergraph, then the two adjacency tensors are permutational similar. Since permutational similar tensors have the same spectra (by Theorem 2.1), we can now obtain an affirmative answer to the above question:

Theorem 2.2: (1). The spectrum (as a multiset) of a hypergraph is independent of the ordering of its vertices.

(2). Isomorphic hypergraphs have the same spectra. \square

Now we consider the second special kind of similarity of tensors: the diagonal similarity. This relation of tensors was first studied in [14] and [15], and was applied to the study of the spectra of nonnegative irreducible tensors. Here we use our tensor product to show that, this relation between tensors is also a special kind of similarity in the sense of Definition 2.3.

Lemma 2.3: Let $D = \text{diag}(d_{11}, \dots, d_{nn})$ be an invertible diagonal matrix of order n . Then we have $D^{-(m-1)}\mathbb{I}D = \mathbb{I}$.

Proof. By using the formula (2.1) we have

$$(D^{-(m-1)}\mathbb{I}D)_{i_1 \dots i_m} = \sum_{j_1, \dots, j_m=1}^n \delta_{j_1 \dots j_m} (D^{-(m-1)})_{i_1 j_1} d_{j_2 i_2} \dots d_{j_m i_m} = \sum_{j=1}^n (D^{-(m-1)})_{i_1 j} d_{j i_2} \dots d_{j i_m} = \delta_{i_1 \dots i_m}$$

\square

Definition 2.4: Let \mathbb{A} and \mathbb{B} be two order m dimension n tensors. We say that \mathbb{A} and \mathbb{B} are diagonal similar, if there exists some invertible diagonal matrix D of order n such that $\mathbb{B} = D^{-(m-1)}\mathbb{A}D$.

Theorem 2.3: If the two order m dimension n tensors \mathbb{A} and \mathbb{B} are diagonal similar, Then $\phi_{\mathbb{B}}(\lambda) = \phi_{\mathbb{A}}(\lambda)$.

Proof. From Lemma 2.3 we see that diagonal similar tensors are similar. \square

Using Theorem 2.3, we can obtain the following result (which is stronger than a related result in [14]) on the symmetry of the spectrum of a nonnegative irreducible tensor.

Theorem 2.4: Let \mathbb{A} be an order m dimension n nonnegative irreducible tensor with the spectral radius $\rho(A)$. Suppose that the cyclic index of \mathbb{A} is k (namely, \mathbb{A} has exactly k distinct eigenvalues with modulus $\rho(A)$). Then $\text{spec}(\mathbb{A})$ (as a multiset) is invariant under a rotation of an angle $\frac{2\pi}{k}$ on the complex plane, and therefore $\phi_{\mathbb{A}}(\lambda) = \lambda^r f(\lambda^k)$ for some nonnegative integer r and some polynomial $f(x)$.

Proof. By [14], there exists an invertible diagonal matrix D , such that

$$\mathbb{A} = e^{-\frac{2\pi i}{k}} D^{-(m-1)} \mathbb{A} D.$$

So from Theorem 2.2 we have $\text{spec}(\mathbb{A}) = e^{-\frac{2\pi i}{k}} \text{spec}(D^{-(m-1)} \mathbb{A} D) = e^{-\frac{2\pi i}{k}} \text{spec}(\mathbb{A})$.

Now suppose that

$$\phi_{\mathbb{A}}(\lambda) = \sum_{j=0}^d a_j \lambda^{d-j}, \quad (\text{where } d \text{ is the degree of } \phi_{\mathbb{A}}(\lambda))$$

Then from the above result we have

$$e^{\frac{2\pi di}{k}} \phi_{\mathbb{A}}(\lambda) = \phi_{\mathbb{A}}(e^{\frac{2\pi i}{k}} \lambda), \quad \text{and thus} \quad e^{\frac{2\pi di}{k}} \sum_{j=0}^d a_j \lambda^{d-j} = \sum_{j=0}^d a_j (e^{\frac{2\pi i}{k}} \lambda)^{d-j}$$

Comparing the coefficients of both sides, we obtain $e^{\frac{2\pi di}{k}} a_j = e^{\frac{2\pi(d-j)i}{k}} a_j$. Thus $a_j = 0$ if j is not a multiple of k . Therefore we have $\phi_{\mathbb{A}}(\lambda) = \lambda^r f(\lambda^k)$. \square

In [14] and [15], Q.Yang and Y.Yang give the following results regarding the diagonal similar tensors. Here we show that the proofs of these results can be simplified by using the associativity of our tensor product.

Theorem 2.5 ([14] and [15]) : (1). Suppose that the two tensors \mathbb{A} and \mathbb{B} are diagonal similar, namely $\mathbb{B} = D^{-(m-1)}\mathbb{A}D$ for some invertible diagonal matrix D . Then x is an eigenvector of \mathbb{B} corresponding to the eigenvalue λ if and only if $y = Dx$ is an eigenvector of \mathbb{A} corresponding to the same eigenvalue λ .

(2). Every nonnegative irreducible tensor \mathbb{A} is diagonal similar to a multiple of some stochastic tensor.

Proof : (1). By the associative law of the tensor product, we have for $y = Dx$ that

$$\mathbb{B}x = \lambda x^{[m-1]} \iff D^{-(m-1)}\mathbb{A}Dx = \lambda x^{[m-1]} \iff \mathbb{A}y = D^{(m-1)}\lambda x^{[m-1]} = \lambda(Dx)^{[m-1]} \iff \mathbb{A}y = \lambda y^{[m-1]}$$

(2). Let u be the positive eigenvector of \mathbb{A} corresponding to the eigenvalue $\rho(\mathbb{A})$. Let e be the column vector of dimension n with all components 1, let D be the invertible diagonal matrix with $u = De$, and let $\mathbb{B} = D^{-(m-1)}\mathbb{A}D$. Then \mathbb{B} is also nonnegative. So from the result (1) and $\mathbb{A}u = \rho(\mathbb{A})u^{[m-1]}$, we have $\mathbb{B}e = \rho(\mathbb{A})e^{[m-1]} = \rho(\mathbb{A})e$, which means that $\frac{1}{\rho(\mathbb{A})}\mathbb{B}$ is a stochastic tensor. \square

Now we consider another relation between tensors: the congruence relation (which is also a generalization of the congruence relation between matrices) and its applications on the \mathbf{E} -eigenvalues and \mathbf{E} -eigenvectors of

tensors. First we recall that in [12, section 6], L.Qi defined a useful relation between tensors (which he called "orthogonal similarity") as follows.

Definition 2.5 (L.Qi, 2005 [12]): Let $P = (p_{ij})$ be a matrix of order n , \mathbb{A} and \mathbb{B} be two order m dimension n tensors. If the elements of \mathbb{A} and \mathbb{B} satisfy the following relation:

$$b_{i_1 \dots i_m} = \sum_{j_1, \dots, j_m=1}^n a_{j_1 \dots j_m} p_{i_1 j_1} p_{i_2 j_2} \dots p_{i_m j_m}.$$

Then it is written as $\mathbb{B} = P^m \mathbb{A}$. In the case when P is a real orthogonal matrix, then \mathbb{A} and \mathbb{B} are called "orthogonal similar".

Remark 2.3: (1) By using the tensor product defined in section 1 and the formula (2.1), we see that the relation $\mathbb{B} = P^m \mathbb{A}$ defined above is actually the same as the tensor relation $\mathbb{B} = P \mathbb{A} P^T$. Since the corresponding relation for matrices is called the "congruence relation", so in this paper we choose to call the two tensors \mathbb{A} and \mathbb{B} to be congruent if $\mathbb{B} = P \mathbb{A} P^T$ for some matrix P . In the case when P is a real orthogonal matrix, then we call \mathbb{A} and \mathbb{B} as "orthogonal congruent".

(2) Notice that when P is a real orthogonal matrix, the relation $P \mathbb{I} P^T = \mathbb{I}$ does not necessarily hold. That is to say, $P \mathbb{A} P^T$ and \mathbb{A} are not necessary similar in the sense of our Definition 2.3. This is another reason why in this paper we choose to call them "orthogonal congruent" instead of "orthogonal similar". \square

It is easy to see that the orthogonal congruence relation is also an equivalent relation between tensors.

In [12, Theorem 7], L.Qi also proved the following result about the **E**-eigenvalues and **E**-eigenvectors of tensors, which can now be simply stated and proved as the following, by using our tensor product and its associative law.

Theorem 2.6 : Suppose that the two tensors \mathbb{A} and \mathbb{B} are "orthogonal congruent", namely $\mathbb{B} = P \mathbb{A} P^T$ for some real orthogonal matrix P . Then x is an **E**-eigenvector of \mathbb{A} corresponding to the **E**-eigenvalue λ if and only if $y = Px$ is an **E**-eigenvector of \mathbb{B} corresponding to the **E**-eigenvalue λ .

Proof : Since $y^T y = x^T x$ for $y = Px$, it is obvious that $x^T x = 1$ if and only if $y^T y = 1$. Also, by the associative law of the tensor product, we have

$$\mathbb{B}y = \lambda y \iff (P \mathbb{A} P^T)(Px) = \lambda(Px) \iff P(\mathbb{A}x) = P(\lambda x) \iff \mathbb{A}x = \lambda x$$

3. The direct product $\mathbb{A} \otimes \mathbb{B}$ of tensors with applications on the spectra of products of hyper-graphs

It is well-known that the direct product of matrices (denoted by \otimes) is a useful concept and tool in matrix theory. It has many applications in various fields, including the applications in the study of the spectra of products of graphs. In this section, we will show that the concept of the direct product of matrices can be generalized to tensors, and can also be applied to the study of the spectrum of the products of hypergraphs.

Definition 3.1: Let \mathbb{A} and \mathbb{B} be two order k tensors with dimension n and m , respectively. Define the direct product $\mathbb{A} \otimes \mathbb{B}$ to be the following tensor of order k and dimension nm (the set of subscripts is taken as

$[n] \times [m]$ in the lexicographic order):

$$(\mathbb{A} \otimes \mathbb{B})_{(i_1, j_1)(i_2, j_2) \dots (i_k, j_k)} = a_{i_1 i_2 \dots i_k} b_{j_1 j_2 \dots j_k}$$

From this definition it is easy to see the following proposition.

Proposition 3.1:

- (1). $(\mathbb{A}_1 + \mathbb{A}_2) \otimes \mathbb{B} = \mathbb{A}_1 \otimes \mathbb{B} + \mathbb{A}_2 \otimes \mathbb{B}$.
- (2). $\mathbb{A} \otimes (\mathbb{B}_1 + \mathbb{B}_2) = \mathbb{A} \otimes \mathbb{B}_1 + \mathbb{A} \otimes \mathbb{B}_2$.
- (3). $(\lambda \mathbb{A}) \otimes \mathbb{B} = \mathbb{A} \otimes (\lambda \mathbb{B}) = \lambda(\mathbb{A} \otimes \mathbb{B}). \quad (\lambda \in \mathbb{C}.)$

The following theorem gives an important relation between the direct product of tensors and the general product of tensors defined in section 1, which is also a generalization of the corresponding relation for matrices.

Theorem 3.1: Let \mathbb{A} and \mathbb{B} be two order $k+1$ tensors with dimension n and m , respectively. Let \mathbb{C} and \mathbb{D} be two order $r+1$ tensors with dimension n and m , respectively. Then we have:

$$(\mathbb{A} \otimes \mathbb{B})(\mathbb{C} \otimes \mathbb{D}) = (\mathbb{A}\mathbb{C}) \otimes (\mathbb{B}\mathbb{D}).$$

Proof. For the simplicity of notation, we write

$$(\alpha, \beta) = ((a_1, b_1), \dots, (a_r, b_r)) \in ([n] \times [m])^r, \quad (\text{if } \alpha = (a_1, \dots, a_r) \in [n]^r, \quad \beta = (b_1, \dots, b_r) \in [m]^r).$$

In the following, we assume $\alpha_i \in [n]^r, \beta_j \in [m]^r, (i, j = 1, \dots, k)$. Then by Definition 1.1 we have

$$\begin{aligned} & ((\mathbb{A} \otimes \mathbb{B})(\mathbb{C} \otimes \mathbb{D}))_{(i, j)(\alpha_1, \beta_1) \dots (\alpha_k, \beta_k)} \\ &= \sum_{(i_1, j_1), \dots, (i_k, j_k) \in [n] \times [m]} (\mathbb{A} \otimes \mathbb{B})_{(i, j)(i_1, j_1) \dots (i_k, j_k)} (\mathbb{C} \otimes \mathbb{D})_{(i_1, j_1)(\alpha_1, \beta_1)} \dots (\mathbb{C} \otimes \mathbb{D})_{(i_k, j_k)(\alpha_k, \beta_k)} \\ &= \sum_{i_1, \dots, i_k=1}^n \sum_{j_1, \dots, j_k=1}^m a_{i i_1 \dots i_k} b_{j j_1 \dots j_k} (c_{i_1 \alpha_1} d_{j_1 \beta_1}) \dots (c_{i_k \alpha_k} d_{j_k \beta_k}) \\ &= \left(\sum_{i_1, \dots, i_k=1}^n a_{i i_1 \dots i_k} c_{i_1 \alpha_1} \dots c_{i_k \alpha_k} \right) \left(\sum_{j_1, \dots, j_k=1}^m b_{j j_1 \dots j_k} d_{j_1 \beta_1} \dots d_{j_k \beta_k} \right) \\ &= (\mathbb{A}\mathbb{C})_{i \alpha_1 \dots \alpha_k} (\mathbb{B}\mathbb{D})_{j \beta_1 \dots \beta_k} \\ &= ((\mathbb{A}\mathbb{C}) \otimes (\mathbb{B}\mathbb{D}))_{(i, j)(\alpha_1, \beta_1) \dots (\alpha_k, \beta_k)} \end{aligned}$$

Thus we have $(\mathbb{A} \otimes \mathbb{B})(\mathbb{C} \otimes \mathbb{D}) = (\mathbb{A}\mathbb{C}) \otimes (\mathbb{B}\mathbb{D})$. □

Now we consider the applications of the direct product of tensors in the study of the spectra of the products of hypergraphs.

Definition 3.2 ([5], The Cartesian product of hypergraphs): Let G and H be two hypergraphs. Define the Cartesian product $G \square H$ of G and H as: $V(G \square H) = V(G) \times V(H)$, and $\{(i_1, j_1), \dots, (i_r, j_r)\} \in E(G \square H)$ if and only if one of the following two conditions holds:

- (1). $i_1 = \dots = i_r$ and $\{j_1, \dots, j_r\} \in E(H)$.
- (2). $j_1 = \dots = j_r$ and $\{i_1, \dots, i_r\} \in E(G)$.

A hypergraph $H = (V, E)$ is called k -uniform if every edge of H is a subset of V with k elements. The adjacency tensor of H (under certain ordering of vertices) is the order k dimension n tensor \mathbb{A}_H with the entries ([5])

$$a_{i_1 i_2 \dots i_k} = \begin{cases} \frac{1}{(k-1)!} & \text{if } \{i_1, i_2, \dots, i_k\} \in E(H) \\ 0 & \text{otherwise} \end{cases}$$

The characteristic polynomial and spectrum of a uniform hypergraph H are that of its adjacency tensor.

Using the direct product of tensors, the adjacency tensor of the Cartesian product of k -uniform hypergraphs can be expressed as in the following theorem.

Theorem 3.2: Let \mathbb{A} and \mathbb{B} be the adjacency tensors of a k -uniform hypergraph G of order n and a k -uniform hypergraph H of order m , respectively. Then the adjacency tensor of $G \square H$ is $\mathbb{A} \otimes \mathbb{I}_m + \mathbb{I}_n \otimes \mathbb{B}$ (where the ordering of the vertices of $G \square H$ is taken to be the lexicographic ordering of the elements of the set $V(G) \times V(H)$).

Proof. Let \mathbb{C} be the adjacency tensor of the hypergraph $G \square H$. Then by definition we can check that:

$$c_{(i_1, j_1) \dots (i_k, j_k)} = \begin{cases} b_{j_1 \dots j_k} & \text{if } i_1 = i_2 = \dots = i_k \\ a_{i_1 \dots i_k} & \text{if } j_1 = j_2 = \dots = j_k \\ 0 & \text{otherwise} \end{cases} \quad (3.1)$$

Notice that all the diagonal entries of \mathbb{A} and \mathbb{B} are zero, so it follows from (3.1) that

$$c_{(i_1, j_1) \dots (i_k, j_k)} = a_{i_1 \dots i_k} \delta_{j_1 \dots j_k} + \delta_{i_1 \dots i_k} b_{j_1 \dots j_k}$$

It is also easy to see that

$$(\mathbb{A} \otimes \mathbb{I}_m + \mathbb{I}_n \otimes \mathbb{B})_{(i_1, j_1) \dots (i_k, j_k)} = a_{i_1 \dots i_k} \delta_{j_1 \dots j_k} + \delta_{i_1 \dots i_k} b_{j_1 \dots j_k}$$

So we have $\mathbb{C} = \mathbb{A} \otimes \mathbb{I}_m + \mathbb{I}_n \otimes \mathbb{B}$. □

Now we consider another kind of product of the hypergraphs.

Definition 3.3 (The direct product of hypergraphs): Let G and H be two hypergraphs. Define the direct product $G \times H$ of G and H as: $V(G \times H) = V(G) \times V(H)$, and $\{(i_1, j_1), \dots, (i_r, j_r)\} \in E(G \times H)$ if and only if $\{i_1, \dots, i_r\} \in E(G)$ and $\{j_1, \dots, j_r\} \in E(H)$.

Using the direct product of tensors, the adjacency tensor of the direct product of k -uniform hypergraphs can be easily obtained as in the following theorem.

Theorem 3.3: Let \mathbb{A} and \mathbb{B} be the adjacency tensors of a k -uniform hypergraph G of order n and a k -uniform hypergraph H of order m , respectively. Then the adjacency tensor of $G \times H$ is $(k-1)!(\mathbb{A} \otimes \mathbb{B})$ (where the ordering of the vertices of $G \times H$ is taken to be the lexicographic ordering of the elements of the set $V(G) \times V(H)$).

Proof. Let \mathbb{C} be the adjacency tensor of the hypergraph $G \times H$. Then by definition we can check that

$$c_{(i_1, j_1) \dots (i_k, j_k)} = \begin{cases} \frac{1}{(k-1)!} & \text{if } a_{i_1 \dots i_k} \text{ and } b_{j_1 \dots j_k} \text{ are both nonzero} \\ 0 & \text{otherwise} \end{cases}$$

Thus it follows that

$$c_{(i_1, j_1) \dots (i_k, j_k)} = (k-1)! a_{i_1 \dots i_k} b_{j_1 \dots j_k} = (k-1)!(\mathbb{A} \otimes \mathbb{B})_{(i_1, j_1) \dots (i_k, j_k)}$$

So we have $\mathbb{C} = (k-1)!(\mathbb{A} \otimes \mathbb{B})$. □

The following theorem gives the relation between the eigenvalue-eigenvectors of the tensors \mathbb{A} and \mathbb{B} and that of $\mathbb{A} \otimes \mathbb{B}$ and $\mathbb{A} \otimes \mathbb{I}_m + \mathbb{I}_n \otimes \mathbb{B}$.

Theorem 3.4: Let \mathbb{A} and \mathbb{B} be two order k tensors with dimension n and m , respectively. Suppose that we have $\mathbb{A}u = \lambda u^{[k-1]}$, and $\mathbb{B}v = \mu v^{[k-1]}$, and we also write $w = u \otimes v$. Then we have:

(1). $(\mathbb{A} \otimes \mathbb{I}_m + \mathbb{I}_n \otimes \mathbb{B})w = (\lambda + \mu)w^{[k-1]}$.

(2). $(\mathbb{A} \otimes \mathbb{B})w = (\lambda\mu)w^{[k-1]}$.

Proof. We have by Proposition 3.1 and Theorem 3.1 that:

(1).

$$\begin{aligned} (\mathbb{A} \otimes \mathbb{I}_m + \mathbb{I}_n \otimes \mathbb{B})w &= (\mathbb{A} \otimes \mathbb{I}_m + \mathbb{I}_n \otimes \mathbb{B})(u \otimes v) = (\mathbb{A}u) \otimes (\mathbb{I}_m v) + (\mathbb{I}_n u) \otimes (\mathbb{B}v) \\ &= (\lambda u^{[k-1]}) \otimes (v^{[k-1]}) + (u^{[k-1]}) \otimes (\mu v^{[k-1]}) \\ &= (\lambda + \mu)(u^{[k-1]} \otimes v^{[k-1]}) = (\lambda + \mu)(u \otimes v)^{[k-1]} = (\lambda + \mu)w^{[k-1]} \end{aligned}$$

(2).

$$\begin{aligned} (\mathbb{A} \otimes \mathbb{B})w &= (\mathbb{A} \otimes \mathbb{B})(u \otimes v) = (\mathbb{A}u) \otimes (\mathbb{B}v) = (\lambda u^{[k-1]}) \otimes (\mu v^{[k-1]}) \\ &= (\lambda\mu)(u^{[k-1]} \otimes v^{[k-1]}) = (\lambda\mu)(u \otimes v)^{[k-1]} = (\lambda\mu)w^{[k-1]}. \end{aligned}$$

□

From Theorem 3.4, we can obtain the following results about the spectra of the hypergraphs $G \square H$ and $G \times H$, where the result on the spectrum of $G \square H$ was obtained by Cooper and Dutle in [5], but by using the direct product of tensors defined here, our Theorem 3.4 actually gives a simplified proof of that result on the spectrum of $G \square H$.

Theorem 3.5: Let G and H be two k -uniform hypergraphs. Let λ be an eigenvalue of G with corresponding eigenvector u , and μ be an eigenvalue of H with corresponding eigenvector v , respectively. Then we have:

(1). ([5]) $\lambda + \mu$ is an eigenvalue of $G \square H$ with corresponding eigenvector $u \otimes v$.

(2). $\lambda\mu$ is an eigenvalue of $G \times H$ with corresponding eigenvector $u \otimes v$. □

4. Applications in the study of nonnegative tensors

In [2] and [4], Chang et al studied the properties of the spectra of nonnegative tensors. They defined the irreducibility of tensors, and the primitivity of nonnegative tensors (also see [11]), and extended the well-known Perron-Frobenius Theorem from the nonnegative irreducible matrices to nonnegative irreducible tensors. They also extended many important properties of primitive matrices to primitive tensors. In [14] and [15], Q. Yang and Y. Yang gave some further results on the generalizations of Perron-Frobenius Theorem, including some results on the symmetry of the spectra of nonnegative irreducible tensors.

In this section, we will give some applications of the tensor product in the study of nonnegative tensors, especially in the study of primitive tensors. Among these applications, we will give a simple characterization of the primitive tensors in terms of the zero patterns of the powers of \mathbb{A} (in the sense of the tensor product

defined in section 1), and give some upper bounds for the primitive degrees of the primitive tensors. We also give some result on the cyclic index of a nonnegative irreducible tensor.

Let $Z(\mathbb{A})$ be the tensor obtained by replacing all the nonzero entries of \mathbb{A} by one. Then $Z(\mathbb{A})$ is called the zero-nonzero pattern of \mathbb{A} (or simply the zero pattern of \mathbb{A}).

First we have the following result on the zero patterns of the products of two nonnegative tensors.

Lemma 4.1: The zero pattern of the product $\mathbb{A}\mathbb{B}$ of two nonnegative tensors \mathbb{A} and \mathbb{B} is uniquely determined by the zero patterns of \mathbb{A} and \mathbb{B} .

Proof. Suppose that $Z(\mathbb{A}_1) = Z(\mathbb{A}_2)$ and $Z(\mathbb{B}_1) = Z(\mathbb{B}_2)$. Then there exist positive numbers c_1, c_2, M_1 and M_2 such that

$$c_1\mathbb{A}_1 \leq \mathbb{A}_2 \leq M_1\mathbb{A}_1; \quad c_2\mathbb{B}_1 \leq \mathbb{B}_2 \leq M_2\mathbb{B}_1$$

From this it follows that $c_1c_2^{m-1}\mathbb{A}_1\mathbb{B}_1 \leq \mathbb{A}_2\mathbb{B}_2 \leq M_1M_2^{m-1}\mathbb{A}_1\mathbb{B}_1$ (where m is the order of \mathbb{A}_1). Thus we have $Z(\mathbb{A}_1\mathbb{B}_1) = Z(\mathbb{A}_2\mathbb{B}_2)$. \square

The following result will be used in Definition 4.1.

Proposition 4.1: Let \mathbb{A} be an order m dimension n nonnegative tensor. Then the following three conditions are equivalent:

- (1). For any $i, j \in [n]$, $a_{ij\dots j} > 0$ holds.
- (2). For any $j \in [n]$, $\mathbb{A}e_j > 0$ holds (where e_j is the j^{th} column of the identity matrix I_n).
- (3). For any nonnegative nonzero vector $x \in \mathbb{R}^n$, $\mathbb{A}x > 0$ holds.

The proof of this proposition is straightforward. \square

Definition 4.1 (Pearson, [10]): A nonnegative tensor \mathbb{A} is called essentially positive, if it satisfies one of the three conditions in Proposition 4.1.

In [4] and [11], Chang et al and Pearson define the primitive tensors as follows.

Definition 4.2 ([4] and [11]): Let \mathbb{A} be a nonnegative tensor with order m and dimension n . Define the map $T_{\mathbb{A}}$ from \mathbb{R}^n to \mathbb{R}^n as: $T_{\mathbb{A}}(x) = (\mathbb{A}x)^{[\frac{1}{m-1}]}$ (here $\mathbb{A}x$ is understood as the tensor product defined in section 1). If there exists some positive integer r such that $T_{\mathbb{A}}^r(x) > 0$ for all nonnegative nonzero vectors $x \in \mathbb{R}^n$, then \mathbb{A} is called primitive, and the smallest such integer r is called the primitive degree of \mathbb{A} , denoted by $\gamma(\mathbb{A})$.

By using the properties of tensor product and Lemma 4.1, we can obtain the following simple characterization for primitive tensors.

Theorem 4.1: A nonnegative tensor \mathbb{A} is primitive if and only if there exists some positive integer r such that \mathbb{A}^r is essentially positive. Furthermore, the smallest such r is the primitive degree of \mathbb{A} .

Proof: For convenience, here we use the notation $\mathbb{A} \simeq \mathbb{B}$ to denote that \mathbb{A} and \mathbb{B} have the same zero pattern.

We first use induction on r to show that for any nonnegative nonzero vectors $x \in \mathbb{R}^n$, $T_{\mathbb{A}}^r(x)$ and $\mathbb{A}^r x$ have the same zero pattern. When $r = 1$, we obviously have $T_{\mathbb{A}}(x) = (\mathbb{A}x)^{[\frac{1}{m-1}]} \simeq \mathbb{A}x$. In general, by the inductive hypothesis, Lemma 4.1 and the associative law of the tensor product we have:

$$T_{\mathbb{A}}^r(x) = T_{\mathbb{A}}(T_{\mathbb{A}}^{r-1}(x)) = (\mathbb{A}T_{\mathbb{A}}^{r-1}(x))^{[\frac{1}{m-1}]} \simeq \mathbb{A}T_{\mathbb{A}}^{r-1}(x) \simeq \mathbb{A}(\mathbb{A}^{r-1}x) = \mathbb{A}^r x.$$

From this we have for any nonzero vector $x \geq 0$ and positive integer r that: $T_{\mathbb{A}}^r(x) > 0 \iff \mathbb{A}^r x > 0$. Thus our result follows directly. \square

Using the similar ideas, the characterization of (nonnegative) irreducible tensors given in [15, Theorem 5.2] can also be simply stated as: A nonnegative tensor \mathbb{A} of dimension n is irreducible if and only if $(\mathbb{A} + \mathbb{I})^{n-1}$ is essentially positive.

The following Lemma 4.2 will be used in the proof of Theorem 4.2.

Lemma 4.2: Let \mathbb{A} be a nonnegative tensor. If \mathbb{A}^r is essentially positive, then \mathbb{A}^{r+1} is also essentially positive.

Proof: By hypothesis, $\mathbb{A}^r x > 0$ for all nonzero vectors $x \geq 0$. It follows that $y = \mathbb{A}x$ is also a nonzero (nonnegative) vector (otherwise we will have $\mathbb{A}^r x = 0$). Therefore we have (for any nonzero vectors $x \geq 0$) that: $\mathbb{A}^{r+1}x = \mathbb{A}^r(\mathbb{A}x) = \mathbb{A}^r y > 0$, so \mathbb{A}^{r+1} is essentially positive. \square

The following theorem shows that, although there are infinitely many primitive tensors of order m and dimension n , their primitive degrees are bounded above and have the following upper bound.

Theorem 4.2 (An upper bound of the primitive degrees): Let \mathbb{A} be a nonnegative primitive tensor with order m and dimension n . Then its primitive degree $\gamma(\mathbb{A}) \leq 2^{n^m}$.

Proof: First we notice that there are altogether 2^{n^m} many different zero patterns for order m dimension n tensors. So among the powers $\mathbb{A}, \dots, \mathbb{A}^{2^{n^m}+1}$, there must be some two powers having the same zero pattern. Thus there exist some integer $k \leq 2^{n^m}$ and some integer $p > 0$ such that $Z(\mathbb{A}^k) = Z(\mathbb{A}^{k+p})$. From this it will follow by Lemma 4.1 that $Z(\mathbb{A}^k) = Z(\mathbb{A}^{k+tp})$ for all nonnegative integers t . On the other hand, by Lemma 4.2 we also know that \mathbb{A}^{k+tp} is essentially positive for all sufficiently large t , so \mathbb{A}^k is essentially positive, which implies by definition that $\gamma(\mathbb{A}) \leq k \leq 2^{n^m}$. \square

Remark 4.1: The upper bound 2^{n^m} can be replaced by the number of different zero patterns of all the primitive tensors of order m and dimension n . Thus the upper bound in Theorem 4.2 is not sharp. \square

In [10], Pearson defined the majorization matrix $M(\mathbb{A})$ of a tensor \mathbb{A} as the following:

$$(M(\mathbb{A}))_{ij} = a_{ij\dots j} \quad (i, j = 1, \dots, n)$$

Pearson ([10]) also proved that if $M(\mathbb{A})$ is an irreducible matrix, then \mathbb{A} is also irreducible.

Also, the spectral radius of \mathbb{A} is denoted by $\rho(\mathbb{A})$, and the cyclic index of a nonnegative irreducible tensor \mathbb{A} is defined to be the number of distinct eigenvalues of \mathbb{A} whose absolute value is $\rho(\mathbb{A})$.

In the following theorem, we use the majorization matrix $M(\mathbb{A})$ to study the cyclic index of a nonnegative tensor \mathbb{A} .

Theorem 4.3: Let \mathbb{A} be a nonnegative tensor with order m and dimension n . If the matrix $M(\mathbb{A})$ is irreducible, then the cyclic index of $M(\mathbb{A})$ is a multiple of the cyclic index of \mathbb{A} .

Proof: By ([10]) we know that \mathbb{A} is also irreducible. Now let k be the cyclic index of \mathbb{A} , and p be the cyclic index of the irreducible matrix $M(\mathbb{A})$. If $k = 1$, the result holds obviously. So we assume that $k \geq 2$.

Let $\lambda = \rho(\mathbb{A})e^{i2\pi/k}$, let $y = (y_1, \dots, y_n)^T$ be an eigenvector of \mathbb{A} (with full support) corresponding to the eigenvalue λ , and write (as in [4])

$$y_j = |y_j|e^{i2\pi\phi_j} \quad (j = 1, \dots, n)$$

Then by [4, Lemma 4.4], we have for any $a_{i_1 \dots i_m} > 0$ that

$$\phi_{i_2} + \dots + \phi_{i_m} - (m-1)\phi_{i_1} - \frac{1}{k} \in \mathbb{Z} \quad (\mathbb{Z} \text{ is the set of all integers}) \quad (4.1)$$

Let $j_1 \dots j_r j_1$ be a (directed) cycle of the associated digraph $D(M(\mathbb{A}))$ ([1]) of the matrix $M(\mathbb{A})$ (with length r), then we have

$$\begin{aligned} (M(\mathbb{A}))_{j_t j_{t+1}} > 0 \quad (t = 1, \dots, r, \text{ mod } r) &\implies a_{j_t j_{t+1} \dots j_{t+1}} > 0 \quad (t = 1, \dots, r, \text{ mod } r) \\ &\implies (m-1)\phi_{j_{t+1}} - (m-1)\phi_{j_t} - \frac{1}{k} \in \mathbb{Z} \quad (t = 1, \dots, r, \text{ mod } r) \end{aligned}$$

Adding the above r relations together, we obtain

$$(m-1) \sum_{t=1}^r (\phi_{j_{t+1}} - \phi_{j_t}) - \frac{r}{k} = -\frac{r}{k} \in \mathbb{Z} \quad (4.2)$$

Thus k divides r . But r is an arbitrary cycle length of the associated digraph of the matrix $M(\mathbb{A})$, so k divides the greatest common divisor of all the cycle lengths of the associated digraph of $M(\mathbb{A})$, which is just p , the cyclic index of the irreducible matrix $M(\mathbb{A})$ ([1]). \square

We also have the following relation between the primitive degrees of \mathbb{A} and $M(\mathbb{A})$.

Corollary 4.1: Let \mathbb{A} be a nonnegative tensor with order m and dimension n . If $M(\mathbb{A})$ is primitive, then \mathbb{A} is also primitive and in this case, we have $\gamma(\mathbb{A}) \leq \gamma(M(\mathbb{A})) \leq (n-1)^2 + 1$.

Proof: $\gamma(M(\mathbb{A})) \leq (n-1)^2 + 1$ is the well-known Wielandt's upper bound for primitive matrices ([16]). So we only need to show that \mathbb{A} is primitive and $\gamma(\mathbb{A}) \leq \gamma(M(\mathbb{A}))$. For this purpose, we use the notation $\mathbb{A} \succsim \mathbb{B}$ to denote (the zero patterns) $Z(\mathbb{A}) \geq Z(\mathbb{B})$. Then we first prove for two nonnegative tensors \mathbb{A} and \mathbb{B} that $M(\mathbb{A}\mathbb{B}) \succsim M(\mathbb{A})M(\mathbb{B})$. For if $(M(\mathbb{A})M(\mathbb{B}))_{ij} > 0$, then there exists $h \in [n]$ such that $(M(\mathbb{A}))_{ih} > 0$ and $(M(\mathbb{B}))_{hj} > 0$. It means that $a_{ih \dots h} > 0$ and $b_{hj \dots j} > 0$, which implies that

$$(M(\mathbb{A}\mathbb{B}))_{ij} = (\mathbb{A}\mathbb{B})_{ij \dots j} = \sum_{i_2, \dots, i_m=1}^n a_{ii_2 \dots i_m} b_{i_2 j \dots j} \dots b_{i_m j \dots j} \geq a_{ih \dots h} (b_{hj \dots j})^{m-1} > 0.$$

Now from $M(\mathbb{A}\mathbb{B}) \succsim M(\mathbb{A})M(\mathbb{B})$ we can see that for any integer $k > 0$, we have $M(\mathbb{A}^k) \succsim (M(\mathbb{A}))^k$. Take $k = \gamma(M(\mathbb{A}))$, then we have $M(\mathbb{A}^k) \succsim (M(\mathbb{A}))^k > 0$, which means that \mathbb{A}^k is essentially positive. So \mathbb{A} is also primitive and $\gamma(\mathbb{A}) \leq k = \gamma(M(\mathbb{A}))$. \square

Now we propose the following conjecture on the primitive degrees.

Conjecture 1: When m is fixed, then there exists some polynomial $f(n)$ on n such that $\gamma(\mathbb{A}) \leq f(n)$ for all nonnegative primitive tensors of order m and dimension n .

In the case of $m = 2$ (\mathbb{A} is a matrix), the Wielandt's upper bound tells us that we can take $f(n) = (n-1)^2 + 1$.

Definition 4.3: Let \mathbb{A} be a nonnegative tensor with order m and dimension n . If there exists some positive integer k such that $\mathbb{A}^k > 0$ is a positive tensor, then \mathbb{A} is called strongly primitive, and the smallest such k is called the strongly primitive degree of \mathbb{A} .

It is obvious that the strong primitivity implies the primitivity. Also, in the matrix case ($m = 2$), it is well-known ([1]) that a nonnegative matrix A is primitive if and only if it is strongly primitive. But the following example shows that these two concepts are not equivalent in the case $m \geq 3$.

Example 4.1: Let \mathbb{A} be a nonnegative tensor of order $m \geq 3$ and dimension n , and with

$$a_{i_1 i_2 \dots i_m} = \begin{cases} 1 & \text{if } i_2 = \dots = i_m \\ 0 & \text{otherwise} \end{cases}$$

Then \mathbb{A} is a primitive tensor, but not a strongly primitive tensor.

Proof: \mathbb{A} is primitive since \mathbb{A} itself is already essentially positive. Now we use induction on k to show that if $i_2 \neq i_3$, then $(\mathbb{A}^k)_{i_1 i_2 i_3 \dots i_r} = 0$ (where $r = (m-1)^k + 1$).

If $k = 1$, the result follows from the definition of \mathbb{A} . In general when $k \geq 2$, let $\mathbb{B} = \mathbb{A}^{k-1} = (b_{j_1 \dots j_t})$, where $t = (m-1)^{k-1} + 1$. Then by the inductive hypothesis, we have $b_{j_1 i_2 i_3 \dots j_t} = 0$ when $i_2 \neq i_3$. Therefore we have

$$(\mathbb{A}^k)_{i_1 i_2 i_3 \dots i_r} = (\mathbb{A}\mathbb{B})_{i_1 i_2 i_3 \dots i_r} = \sum_{j_1, \dots, j_{m-1}=1}^n a_{i_1 j_1 \dots j_{m-1}} b_{j_1 i_2 i_3 \dots i_t} b_{j_2 i_{t+1} \dots i_{2t-1}} \dots b_{j_{m-1} i_{r-t+2} \dots i_r} = 0 \quad (i_2 \neq i_3)$$

Thus \mathbb{A}^k is not a positive tensor for all integers k , so \mathbb{A} is not strongly primitive. \square

References

References

- [1] R.A.BRUALDI AND H.J.RYSER, Combinatorial matrix theory. *Cambridge University Press*, 1991.
- [2] K.C.CHANG, K.PEARSON, AND T.ZHANG, Perron-Frobenius theorem for nonnegative tensors. *Commun. Math. Sci.* 6(2008), 507-520.
- [3] K.C.CHANG, K.PEARSON, AND T.ZHANG, On eigenvalues of real symmetric tensors. *J. Math. Anal. Appl.* 350(2009), 416-422.
- [4] K.C.CHANG, K.PEARSON, AND T.ZHANG, Primitivity, the convergence of the NQZ method, and the largest eigenvalue for nonnegative tensors. *SIAM J. Matrix Anal. Appl.*, 32(2011), 806-819.
- [5] J.COOPER AND A.DUTLE, Spectra of uniform hypergraphs. *Linear Alg. Appl.*, 436(2012), 3268-3292.
- [6] D.COX, J.LITTLE AND D.O'SHEA, Using algebraic geometry. *Springer-Verlag, New York*, 1998.
- [7] I.GELFAND, M.KAPRANOV AND A.ZELEVINSKY, Discriminants, Resultants and Multidimensional Determinants *Birkhauser Boston*, 1994.
- [8] L.H.LIM, Singular values and eigenvalues of tensors, a variational approach, in *Proceedings 1st IEEE international workshop on computational advances of multitensor adaptive processing (2005)*, 129-132.

- [9] M. NG, L. QI AND G. ZHOU, Finding the largest eigenvalue of a nonnegative tensor. *SIAM J. Matrix Anal. Appl.*, 31(2009), 1090-1099.
- [10] K. PEARSON, *Essentially positive tensors*. *Int. J. Algebra*, 4(2010), 421-427.
- [11] K. PEARSON, Primitive tensors and convergence of an iterative process for the eigenvalues of a primitive tensor. *arXiv: 1004.2423v1*, 2010.
- [12] L. QI, Eigenvalues of a real supersymmetric tensor. *Symbolic Comput.*, 40(2005), 1302-1324.
- [13] L. QI, W. SUN AND Y. WANG, Numerical multilinear algebra and its applications. *Front. Math. China*, 2(2007), 501-526
- [14] Y. YANG AND Q. YANG, Further results for Perron-Frobenius theorem for nonnegative tensors. *SIAM J. Matrix Anal. Appl.*, 31(2010), 2517-2530.
- [15] Y. YANG AND Q. YANG, Further results for Perron-Frobenius theorem for nonnegative tensors II. *SIAM J. Matrix Anal. Appl.*, 32(2011), 1236-1250.
- [16] H. WIELANDT, Unzerlegbare, nicht negative matrizen. *Math. Zeit*, 52(1950), 642-648.